

UDC 575:633.15
DOI: 10.2298/GENSR1401049F
Original scientific paper

DETERMINATION RELEVANT BREEDING CRITERIA BY THE PATH AND FACTOR ANALYSIS IN MAIZE

Milomir FILIPOVIĆ¹, Milosav BABIĆ¹, Nenad DELIĆ¹, Goran BEKAVAC²,
Vojka BABIĆ^{*1}

¹Maize Research Institute, Zemun Polje, Belgrade, Serbia

²Institute of Field and Vegetable Crops, Novi Sad, Serbia

Filipovic M., M. Babić, N. Delić, G.Bekavac and V. Babić (2014):
*Determination relevant breeding criteria by the path and factor analysis in
maize-*. Genetika, Vol 46, No. 1, 49-58.

In the process of plant breeding the application of relevant breeding criteria is very important. The Path analysis is broadly applied with the aim to define yield components that mostly determine the yield and that can be used as quality breeding criteria. However, the significance of revealed relationships between yield and yield components can be affected by various factors, such as diverse genetic material that is observed, traits included into analysis, environments in which the material is observed, as well as, the applied statistic approach to determine the nature of the relationships itself.

The interrelationships of yield and yield components of 15 commercial maize hybrids were observed using the Path and factor analyses. According to results of Path analysis, plant height, ear diameter and grain moisture had highly significant genetic and phenotypic direct effects on grain yield. At the same time, factor analysis points out significant effects of two factors on grain yield. Factor 1 was mostly determined by ear length and number of kernels per row, while grain moisture content, ear and cob diameter mostly determined Factor 2.

Key words: common structure, maize, Path coefficients, yield components

INTRODUCTION

The yield, by its nature, is a multidimensional trait that encompasses several different properties and is affected by numerous factors. The application of relevant breeding criteria is very important in the process of breeding and selection of superior genotypes. Therefore,

Corresponding author: Babić Vojka, Maize Research Institute, Zemun Polje, Slobodana Bajića
1, 11185 Belgrade, Serbia; Tel: +38111 37 56 704; Fax: +38111 37 56 707; e-mail:
vbabic@mrizp.rs

knowledge of the relationships between yield and yield components can significantly improve efficiency of breeding programmes. Besides degree of compatibility between two traits obtained by calculating simple correlations it is also desirable to determine relations between independent variables and the trait observed as a dependent variable. The answer to this question may be given by a form of the regression analysis, the method of Path coefficients (Path analysis), which allows perception of direct and indirect effects as well as the proportion of the joint action (determination) of independent variables (x_1, x_2, \dots, x_n) on the dependent variable (y) (WRIGHT, 1934).

BEKAVAC *et al.* (2002) have observed relationships between the stay green trait and some vegetative traits in two broad genetic base maize populations applying the Path analysis and have found out that the stay green had been to the greatest extent correlated with the leaf and stalk water content. The most consistent correlations, both genetic and phenotypic, have been expressed between the yield and the plant height (DRINIĆ *et al.*, 1992). Same authors points out the negative direct effect of the ear length on grain yield, but indirect effect of this trait via plant height has been greater and positive. MUHAMMAD *et al.* (2003) have studied phenotypic and genetic correlations and the direct and indirect effect of various factors on maize yields under common and water stress conditions. Since yield per plant predominantly depends on the process of photosynthesis, selection based on traits that can enhance this process can improve the yield more significantly.

However, when interpreting the results of the Path analysis difficulties in the interpretation of the actual participation of each variable when they are complexly interconnected can arise (SAMONTE *et al.*, 1998; HAIR *et al.*, 1995). SAMONTE *et al.* (1998) have adapted the Path analysis to define interrelationships between yield and yield-related traits in rice by organising and analysing various predictor variables and the first-, second- and third-order variables. MOHAMMADI *et al.*, (2003) considered that this sequential Path mode was advantageous over the conventional Path model in analysing yields and yield components, as special attention was paid to collinearity of various predictor variables. Thus, their actual value in estimating variable-related traits is better predicted.

PETERSON and PFEIFFER (1989) stated that the factor analysis can be applied with the aim to better understand the background of the yield structure as well as the relationship between the yield and yield components. The analysis applied to yield components can indicate basic uncorrelated factors that determine the yield to the greatest extent. BABIĆ *et al.*, (2011) have determined that different planting densities on same maize hybrids studied in the same locations over the same years have resulted in defining different yield components that defined the yield to the greatest extent. These authors have concluded that greater attention should be paid to plant robustness (height), ear length and number of kernels per row for lower planting densities. On the other hand, in the selection process of hybrids intended for increased planting densities, attention should be paid to ear and cob diameters as well as to grain moisture.

The present study encompasses the comparison of results obtained by both analysis, Path and factor, used to reveal interrelationships of yield and yield components. Defining the components that mostly affect the yield formation points out to their significance as breeding criteria in the process of selection for yield improvement.

MATERIALS AND METHODS

The following 15 commercial maize hybrids of different FAO maturity groups (400-700) were included into trials: ZP-42a, ZP-480, NS-501, ZP-533, ZP-599, ZP-570, ZP-580, ZP-633, ZP-677, ZP-704, ZP-701, ZP-753, ZP-732, ZP-735 and NS-640. The four-replicate trials were set up according to the randomised-complete block design in 24 environments. Elementary plot size was 7.392m². The following traits were analysed: Grain yield (kg ha⁻¹), plant height (cm), par height (cm), number of kernel rows, number of kernels per row, ear length (cm), ear diameter (cm), cob diameter (mm) and grain moisture content at harvest (%).

The total number of plants and lodged and broken plants were estimated immediately prior to harvest. The ear yield per each plot was measured at harvest. The cob weight and the moisture content were determined in the average sample of five ears per each hybrid in each of four replicates, while 10 ears per each hybrid per replication were analysed to estimate yield components. Furthermore, 10 plants per each hybrid per replication were used to analyse the plant and ear height.

Interrelationships of observed traits were estimated by the Path analysis (Method of path coefficients, WRIGHT, 1934) in Microsoft Office Excel, 2003 in order to define direct and indirect effects of certain yield components on the yield. The factor analysis (SPSS 15.0 for Windows Evaluation Version, program) of yield components was done with the aim to determine their common structure and to which extent certain yield components define the yield.

RESULTS AND DISCUSSION

As for simple correlations of observed yield components and yields, the both correlations, genetic and phenotypic, were statistically highly significant ($p < 0.01$) for the following traits: Plant height, ear height, ear diameter and moisture content. Such correlations point out that these traits affect the yield formation to a greater extent than remaining observed traits. As correlations are positive it can be concluded that the enhancement of these yield components in the process of selection will result in higher yields.

A method of Path coefficients (WRIGHT, 1934) is a form of the regression analysis that provides an introspective of direct and indirect effects, as well as, of a proportion of co-effects (determinations) of the independent variables (x_1, x_2, \dots, x_n) on the dependent variable (y). The upper number in the Path diagram is the value of genetic correlations, while the lower number is the value of phenotypic correlations and Path coefficients (Figure 1). The Path coefficient indicates the direct effect of the independent variable on the dependent variable. The indirect effects of an independent variable over another independent variable on the dependent variable are gained as a product of coefficients of simple correlation of two independent variables and their individual direct effect.

The values of simple correlations for observed maize hybrids that point out to the degree of compatibility between two traits and values of direct effects of Path coefficients differ. The test of significance of Path coefficients is calculated by standard error with the application of the F test. Both, genetic and phenotypic, direct effects on the yield were highly statistically significant ($p < 0.01$) for the following traits: Plant height, ear diameter and the grain moisture content. Direct phenotypic effects on the yield were highly statistically significant ($p < 0.01$) for the following traits: Ear height, kernel row number, while they were statistically significant ($p < 0.05$) for the number of kernels per row. Genetic and phenotypic direct effects for the ear length, as well as phenotypic direct effects for the ear height, kernel row number and the number

of kernels per row were not statistically significant (Figure 1). Values of direct effects of yield components on the yield are positive except for the grain moisture content. Results on the grain moisture content differ to the greatest extent from the results of simple correlations. The Path analysis revealed that the increase in grain moisture did not result in the direct increase of yields. On the contrary, it decreased yields. However, at the same time, the moisture content positively affected the yield via the ear diameter (Table 1). This means that the yield decreases by the increase in grain moisture, but the increase of the ear diameter most often results in the increase in grain moisture, which leads to significant yield increases. Understanding genetic and non-genetic factors that effects grain quality and grain yield is crucial in developing new cultivars (PRŽULJ *et al.*, 2013).

Relationships between component traits can be different, from absent of significant relationship till almost functional relationship and interaction in compensatory patterns (GARCIA DEL MORAL *et al.*, 2003). Due to cause-effect relationships between variables in a complex system simple correlations analysis may not provide reliable information of the importance of each component trait in determining mega-trait. Path analysis is one of the reliable statistical techniques which partitions correlation coefficients into direct and indirect effects (GUILLEN-PORTAL *et al.*, 2006).

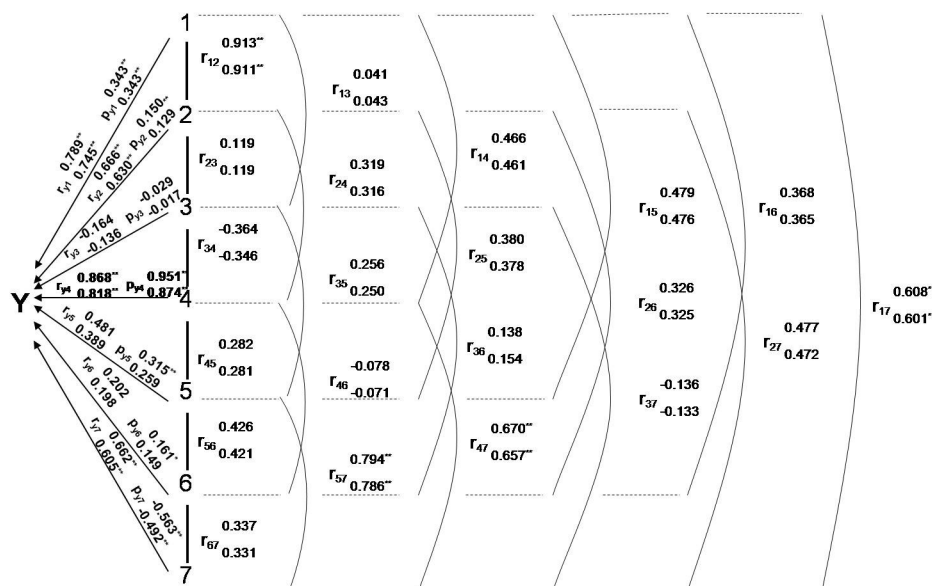


Figure 1. Path diagram for grain yield (Y) and plant height (1), ear height (2), ear length (3), ear diameter (4), kernel row number (5), number of kernels per row (6), grain moisture content (7)

The test of significance cannot be applied to indirect effects, but they are considered significant if their value is higher than the value of direct effects. Although, direct effects of the ear length on the yield are not statistically significant and are negative, the indirect effects of this trait via the kernel row number and the grain moisture content on the yield are significant and positive. This means that the increase in the ear length results in the yield increase via the kernel row number and the grain moisture content. This trait also adversely affects yields via the ear diameter. As these two traits are negatively correlated it means that the increase in the ear length will result in the reduction in both, the ear diameter and yields. Furthermore, the plant height has significant indirect genetic and phenotypic effects on grain yields via the ear diameter. It means that the higher plant height is the greater ear diameter is (the traits are positively correlated), which results in higher yields. The ear height has significant indirect effects on the yield via the plant height and the ear diameter. This means that the increase of the total plant height, which is often followed by the increase of the ear height and ear diameter, has a positive effect on the yield increase. However, the ear height significantly negatively affects the yield via grain moisture. Therefore, the higher ear height is, the greater moisture is, most probably due to the increase of the plant height and the ear diameter, which adversely affects the yield.

Although the kernel row number has a highly statistically significant positive genetic effect on the yield, at the same time it has a significantly negative indirect effect via grain moisture. This means that the higher kernel row number is the higher yield is, but at the same time, the higher grain moisture content is, and it adversely affects the yield. Regarding the kernel row number, selection should be aimed at the increase of this number with the simultaneous reduction of grain moisture.

The coefficient of multiple determination (R^2) pointing out to the common effects of observed traits on the yield was calculated on the basis of values of genetic and phenotypic correlations and values of direct effects on the yield. The coefficient of genetic determination ($R^2_{y1234567}$) was 0.992, meaning that the genetic effect on the yield amounted to 99.2% and was statistically very significant ($p < 0.01$). The coefficient of phenotypic determination was also statistically very significant ($p < 0.01$) and amounted to 0.887, meaning that the phenotypic effect on the yield was 88.7%, i.e. the observed traits determine, to a great extent, the yield.

Speaking of maize hybrid ideotype in context of these results, includes higher plant accompanied with a larger ear diameter and increased number of kernel rows per ear, as well as decrease in grain moisture content and ear position. According to these results, it can be concluded that the ear length has no direct positive effects on the yield and is not correlated to the yield. However, the ear length is negatively correlated to the ear diameter and at the same time, it has a negative indirect effect via the ear diameter on the yield. It means that the greater ear length is the smaller ear diameter is and therefore the lower yield is. In addition, the smaller ear length means greater ear diameter, which has a positive effect on the yield. At the same time, the grain moisture content is greater and the yield is lower. Undoubtedly, these results suggest the need to find a certain balance between the ear length and the ear diameter.

The basic assumption in the application of multiple regression is that the yield components are independent of each other. Actually, yield components are in complex interrelationships that result in multicollinearity. Highly statistically significant ($p < 0.01$) genetic and phenotypic correlations were obtained between the following pairs of traits: plant height and ear height; ear diameter and grain moisture; kernel row number and grain moisture, while genetic and phenotypic correlations between plant height and grain moisture were statistically

significant (Figure 1). Such results indicate that these traits are strongly correlated between themselves and therefore attention has to be paid to this fact in the process of breeding. Multicollinearity of yield components can mask the actual significance of a certain trait or on the other hand in can overvalue some other trait. The conventional Path analysis does not take into account multicollinearity of independent variables and therefore difficulties can arise in the interpretation of results of the actual participation of each variable. Thus, the factor analysis was applied in order to understand the common structure of observed yield components as well as their participation in the formation of principal factors that define the yield.

Table . Direct (diagonals) and indirect effects of yield components (P–phenotypic; G– genotypic)

Trait		PH	EH	EL	ED	KRN	NKR	GM
PH	G	0.343**	0.313	0.014	0.160	0.164	0.126	0.208
	P	0.343**	0.312	0.015	0.158	0.163	0.125	0.206
EH	G	0.137	0.150**	0.018	0.048	0.057	0.049	0.071
	P	0.118	0.129	0.015	0.041	0.049	0.042	0.061
EL	G	-0.001	-0.003	-0.029	0.011	-0.007	-0.004	0.004
	P	-0.001	-0.002	-0.017	0.006	-0.004	-0.003	0.002
ED	G	0.443	0.303	-0.346	0.951**	0.268	-0.074	0.637
	P	0.403	0.276	-0.302	0.874**	0.246	-0.062	0.574
KRN	G	0.151	0.120	0.081	0.089	0.315**	0.134	0.250
	P	0.123	0.098	0.065	0.073	0.259	0.109	0.204
NKR	G	0.059	0.052	0.022	-0.013	0.069	0.161*	0.054
	P	0.054	0.048	0.023	-0.011	0.063	0.149	0.049
GM	G	-0.342	-0.268	0.077	-0.377	-0.447	-0.190	-0.563**
	P	-0.295	-0.232	0.065	-0.323	-0.386	-0.163	-0.491**

PH-Plant height, EH- Ear height, EL-Ear length, ED-Ear diameter, KRN-Kernel row number, NKR-Number of kernels per row, GM-Grain moisture.

Prior to application of the factor analysis, it is necessary to test adequacy of a data set for a given analysis. Values of Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy point out to the participation of the variance of observed variables that can be explained by extracted factors. The values ranging from 0.5 to 1 point out that the factor analysis can be useful for the given data. The value of the KMO Measure of Sampling Adequacy for this set of variables is 0.749. Bartlett's test of sphericity tests the homogeneity of the correlation matrix (identity matrix) and answers to the question of whether the observed variables may be uncorrelated and in such a case inadequate for establishing the common structure, i.e. all diagonal elements are 1 and all off-diagonal elements are 0, implying that all of the variables are uncorrelated. If the test is not significant, as in the case of the given data, the variables share the common structure and the factor analysis can be applied (Table 2).

Two main factors were extracted out of eight observed traits and they encompassed 60.509% of total variability. In such a way, a problem of multicollinearity with a great number of

yield components was minimised to two main factors that were not correlated. Each variable has a portion of variability that shares with remaining variables but it also has a uniqueness that cannot be encompassed by factors. A total of 9.116% (69.625%-60.509%) of variability of yield components for the given data set cannot be encompassed by the factor model (Table 3).

Table 2. KMO and Bartlett's test for assessing the adequacy of the data for the factor analysis

Type of test	Indicators of significance	
KMO Measure of Sampling Adequacy		0.75
Bartlett's test of sphericity	χ^2	8429.56
	df	28
	significance	0.00

Table 3. Portion of the variance of observed traits of maize hybrids encompassed with the first two factor axis

Factor	Initial eigenvalues			Extraction sums of squared loadings		
	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %
1	4.149	51.868	51.868	3.817	47.710	47.710
2	1.421	17.757	69.625	1.024	12.799	60.509
3	0.883	11.038	80.663			
4	0.725	9.068	89.731			
5	0.354	4.423	94.154			
6	0.261	3.258	97.412			
7	0.155	1.935	99.347			
8	0.052	0.653	100.000			

Extraction Method: Principal Axis Factoring.

Due to the existence of the common structure of yield components, variability of each component individually can be partially explained with remaining variables. For instance, the plant height can be explained with 91.3% of other yield components, while the explanation of the kernel row number with other traits is insignificant (31.3%). Variability of all observed traits was to a satisfactory extent encompassed by two extracted factors (56.6 to 79.3% of their variance) except for the kernel row number whose variability was encompassed with only 18.0%, which explicitly indicate that this trait does not share a common structure with remaining yield components (Table 4). Behind eight observed yield components there are two main factors determining the yield of studied maize hybrids. The yield is statistically significantly correlated ($p < 0.01$) with the first two factors (0.632 and 0.400, respectively).

Since two main factors are extracted, the key issue is to what extent certain yield components are included into their formation. In order to easily interpret correlations between

yield components and extracted factors, the rotation of factor axes was done after the method Varimax with Kaiser Normalisation. It is noticeable that the first factor encompassing 47.710% (Table 3) of the variability is in the greatest correlation with the ear length and the number of kernels per row (0.750 and 0.776, respectively). The second factor encompassing 12.799% of the variability is in the greatest correlation with the ear diameter, cob diameter and grain moisture (0.710, 0.756 and 0.742, respectively). Furthermore, it can be seen that the plant height (0.657 and 0.601) and the ear height (0.655 and 0.420) are approximately equally correlated with both extracted factor axes (Table 4).

Table 4. Common structure of yield components and their correlation with rotated factor axes

Yield components	Common structure of yield components		Correlations between yield components and rotated factor axes	
	Initial	Extracted	Factor 1	Factor 2
Plant height	0.913	0.793	0.657	0.601
Ear height	0.880	0.606	0.655	0.420
Ear length	0.610	0.596	0.750	0.185
Ear diameter	0.781	0.777	0.522	0.710
Cob diameter	0.713	0.713	0.376	0.756
Kernel row number	0.313	0.180	0.330	0.267
Number of kernels per row	0.510	0.610	0.776	-0.088
Grain moisture	0.459	0.566	-0.127	0.742

As the first axis (which includes the highest percentage of variance) is highly significantly correlated with the yield, and at the same time largely defined by the ear length and the number of kernels per row, it can be concluded that the effect of these traits on the yield of observed hybrids was the greatest. Furthermore, the second axis is also correlated with obtained yields and it is determined with the ear and cob diameters and grain moisture. This analysis points out that the kernel row number did not share the common structure with remaining observed parameters, which indicates that this trait is independent of the others and that as such can be considered in the process of breeding.

CONCLUSION

By the identification of components that affect yield to the greatest extent, their importance, as breeding criteria in the process of breeding and selection of superior genotypes, is emphasised. Various factors can affect the relationships among yield and yields components, including different germplasm that is included into the study, environmental factors variations, but also the applied statistical analysis itself that is used with the aim to reveal the existing links among observed traits. Different breeding criteria should be applied in the selection process of hybrids for diverse growing conditions. Therefore, it is the most desirable to perform selection under conditions under which hybrids will be grown (FILIPOVIĆ *et al.*, 2013).

ACKNOWLEDGEMENT

Part of this work was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia through the Project TR 31068.

Received November 27th, 2013

Accepted March 14th, 2014

REFERENCES

- BABIĆ, V., M. BABIĆ, V. ANDJELKOVIĆ, V. PERIĆ, and M. SREBRIĆ (2011): Defining primary selection criteria by the factor model. Book of Proceedings of the 46th Croatian and 6th International Symposium on Agriculture, February 14-18, 2011, Opatija, Croatia, pp. 367-370 (CD issue).
- BEKAVAC, G., M. STOJAKOVIĆ, M. IVANOVIĆ, DJ. JOCKOVIĆ, N. VASIĆ, B. PURAR, J. BOČANSKI, and A. NASTASIĆ (2002): Relationships of stay green in maize. *Genetika* 34 (1): 33-40.
- DRINIĆ, G., L. KOJIĆ, M. BABIĆ and J. VANČETOVIĆ (1992): Multiple regression analysis for grain yield and four agronomically important traits of maize (*Zea mays* L.). *Genetika* 24 (3): 195-200.
- FILIPOVIĆ, M., J. SRDIĆ, M. SIMIĆ, Ž. VIDENOVIĆ, Č. RADENOVIĆ, Z. DUMANOVIĆ, and Ž. JOVANOVIĆ (2013): Potential of early maturity flint and dent maize hybrids at higher altitudes. *Romanian Agricultural Research*, 30: 117-124.
- GARCÍA DEL MORAL, L.F., Y. RHARRABTI, D. VILLEGAS, C. ROYO (2003): Evaluation of Grain Yield and Its Components in Durum Wheat under Mediterranean Conditions: An Ontogenic Approach. *Agronomy Journal* 95: 266-274.
- GUILLEN-PORTAL, F.R., R.N. STOUGGARD, Q. XUE, K.M. ESKRIDE (2006): Compensatory mechanisms associated with the effect of spring wheat seed size on wild oat competition. *Crop Science* 46: 935-945.
- HAIR, J.R., R.E. ANDERSON, R.L. TATHAM, and W.C. BLACK (1995): *Multivariate data analysis with readings*. Prentice Hall, Englewood, NY, USA.
- MUHAMMAD, T., I.H. MUHAMMAD, A.A. KIANI, and N. KAMAL (2003): Phenotypic stability for grain yield in maize genotypes under varied rainfield environments. *Asian Journal of Plant Science* 2 (1): 80-82.
- MOHAMMADI, S.A., B.M. PRASANNA, and N.N. SINGH (2003): Sequential path model for determining interrelationships among grain yield and related characters in maize. *Crop Sci.* 43: 1690-1697.
- PETERSON, C.J. and W.H. PFEIFFER (1989): International winter wheat evaluation: Relationships among test sites based on cultivar performance. *Crop Sci.*, 29: 276-282.
- PRŽULJ, N., V. MOMČILOVIĆ and J. CRNOBARAC (2013). Path coefficient analysis of quality of two-row spring barley. *Genetika*, Vol 45 (1): 21-30.
- SAMONTE, S.O.P.B., L.T. WILSON, and A.M. MCCLUNG (1998): Path analysis of yield and yield-related traits of fifteen diverse rice genotypes. *Crop Sci.* 38: 1130-1136.
- WRIGHT, S. (1934): The Methods of Path Coefficients. *Annals of Mathematical Statistics*, 5: 161-215.

UTVRDJIVANJE RELEVANTNIH OPLEMENJIVAČKIH KRITERIJUMA KOD KUKURUZA PATH I FAKTORSKOM ANALIZOM

Milomir FILIPOVIĆ¹, Milosav BABIĆ¹, Nenad DELIĆ¹, Goran BEKAVAC²,
Vojka BABIĆ¹

¹Institut za kukuruz, Zemun Polje, Beograd, Srbija

²Institut za ratarstvo i povrtarstvo, Novi Sad, Srbija

Izvod

Primena relevantnih oplemenjivačkih kriterijuma u oplemenjivanju biljaka je veoma važna. *Path* analiza se široko primenjuje sa ciljem da se definišu komponente prinosa koje u najvećoj meri određuju prinos i koje se mogu uzeti kao pouzdani kriterijumi u procesu oplemenjivanja. Značaj utvrđenih odnosa između prinosa i komponenti prinosa može biti pod uticajem različitih faktora kao što su razlike u genetici materijala koji se ispituje, osobine koje su uključene u analizu, različite spoljne sredine u kojima se vrše ispitivanja kao i sam statistički pristup koji se primenjuje da se utvrdi priroda međusobnih odnosa osobina koje se ispituju. Međusobni odnosi prinosa i komponenti prinosa 15 komercijalnih hibrida kukuruza ispitivani su pomoću *Path* i faktorske analize. Na osnovu rezultata *Path* analize utvrđeno je da visina biljke, prečnik klipa i sadržaj vlage u znu imaju najveći uticaj na prinos. U isto vreme faktorskom analizom izdvojena su dva značajna nekorelisana faktora koja definišu prinos. Factor 1 je pretežno definisan dužinom klipa i brojem zrna u redu, dok su vlaga u zrnu, prečnik klipa i oklaska u najvećoj meri definisali Faktor 2.

Primljeno 27. XI 2013.

Odobreno 14. III. 2014.